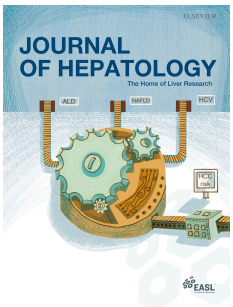
rs641738C>T near MBOAT7 is associated with liver fat, ALT and fibrosis in NAFLD: A meta-analysis

Teo, Kevin ; Abeysekera, Kushala W M ; et al ; Stickel, Felix

Abstract: Background aims: A common genetic variant near MBOAT7 (rs641738C>T) has been previously associated with hepatic fat and advanced histology in non-alcoholic fatty liver disease (NAFLD), however, these findings have not been consistently replicated in the literature. We aimed to establish whether rs641738C>T is a risk factor across the spectrum of NAFLD and characterize its role in the regulation of related metabolic phenotypes through meta-analysis. Methods: We performed meta-analysis of studies with data on the association between rs641738C>T genotype and: liver fat, NAFLD histology, and serum ALT, lipids, or insulin. These included directly genotyped studies and population-level data from genome-wide association studies (GWAS). We performed random effects meta-analysis using recessive, additive, and dominant genetic models. Results: Data from 1,066,175 participants (9,688 with liver biopsies) across 42 studies were included in the meta-analysis. rs641738C>T was associated with higher liver fat on CT/MRI (+0.03 standard deviations [95% CI: 0.02 - 0.05], pz=4.8x10-5) and diagnosis of NAFLD (OR 1.17 [95% CI 1.05 - 1.3], pz=0.003) in Caucasian adults. The variant was also positively associated with presence of advanced fibrosis (OR 1.22 [95% CI: 1.03 - 1.45], pz=0.021) in Caucasian adults using a recessive model of inheritance (CC+CT vs. TT). Meta-analysis of data from previous GWAS found the variant to be associated with higher ALT (pz=0.002) and lower serum triglycerides (pz=1.5x10-4). rs641738C>T was not associated with fasting insulin and no effect was observed in children with NAFLD. Conclusion: Our study validates rs641738C>T near MBOAT7 as a risk factor for the presence and severity of NAFLD in individuals of European descent.

DOI: https://doi.org/10.1016/j.jhep.2020.08.027

Journal Pre-proof



rs641738C>T near MBOAT7 is associated with liver fat, ALT, and fibrosis in NAFLD: a meta-analysis

Kevin Teo, Kushala W.M. Abeysekera, Leon Adams, Elmar Aigner, Quentin M. Anstee, Jesus M. Banales, Rajarshi Banerjee, Priyadarshi Basu, Thomas Berg, Pallav Bhatnagar, Stephan Buch, Ali Canbay, Sonia Caprio, Ankita Chatterjee, Yii-Der Ida Chen, Abhijit Chowdhury, Ann K. Daly, Christian Datz, Dana de Gracia Hahn, Johanna K. DiStefano, Jiawen Dong, Amedine Duret, EU-PNAFLD Investigators, Connor Emdin, Madison Fairey, Glenn S. Gerhard, GOLD Consortium, Xiuqing Guo, Jochen Hampe, Matthew Hickman, Lena Heintz, Christian Hudert, Harriet Hunter, Matt Kelly, Julia Kozlitina, Marcin Krawczyk, Frank Lammert, Claudia Langenberg, Joel Lavine, Lin Li, Hong Kai Lim, Rohit Loomba, Panu K. Luukkonen, Phillip E. Melton, Trevor A. Mori, Nicholette D. Palmer, Constantinos A. Parisinos, Sreekumar G. Pillai, Faiza Qayyum, Matthias C. Reichert, Stefano Romeo, Jerome I. Rotter, Yu Ri Im, Nicola Santoro, Clemens Schafmayer, Elizabeth K. Speliotes, Stefan Stender, Felix Stickel, Christopher D. Still, Pavel Strnad, Kent D. Taylor, Anne Tybjærg-Hansen, Giuseppina Rosaria Umano, Mrudula Utukuri, Luca Valenti, Lynne E. Wagenknecht, Nicholas J. Wareham, Richard M. Watanabe, Julia Wattacheril, Hanieh Yaghootkar, Hannele Yki-Järvinen, Kendra A. Young, Jake P. Mann

PII: S0168-8278(20)33598-4

DOI: https://doi.org/10.1016/j.jhep.2020.08.027

Reference: JHEPAT 7920

To appear in: *Journal of Hepatology*

Received Date: 21 April 2020

Revised Date: 29 July 2020

Accepted Date: 20 August 2020

Please cite this article as: Teo K, Abeysekera KWM, Adams L, Aigner E, Anstee QM, Banales JM, Banerjee R, Basu P, Berg T, Bhatnagar P, Buch S, Canbay A, Caprio S, Chatterjee A, Ida Chen YD, Chowdhury A, Daly AK, Datz C, de Gracia Hahn D, DiStefano JK, Dong J, Duret A, EU-PNAFLD Investigators, Emdin C, Fairey M, Gerhard GS, GOLD Consortium, Guo X, Hampe J, Hickman M, Heintz L, Hudert C, Hunter H, Kelly M, Kozlitina J, Krawczyk M, Lammert F, Langenberg C, Lavine J, Li L, Lim HK, Loomba R, Luukkonen PK, Melton PE, Mori TA, Palmer ND, Parisinos CA, Pillai SG, Qayyum F,

Reichert MC, Romeo S, Rotter JI, Im YR, Santoro N, Schafmayer C, Speliotes EK, Stender S, Stickel F, Still CD, Strnad P, Taylor KD, Tybjærg-Hansen A, Umano GR, Utukuri M, Valenti L, Wagenknecht LE, Wareham NJ, Watanabe RM, Wattacheril J, Yaghootkar H, Yki-Järvinen H, Young KA, Mann JP, rs641738C>T near MBOAT7 is associated with liver fat, ALT, and fibrosis in NAFLD: a meta-analysis, *Journal of Hepatology* (2020), doi: https://doi.org/10.1016/j.jhep.2020.08.027.

MBOAT7

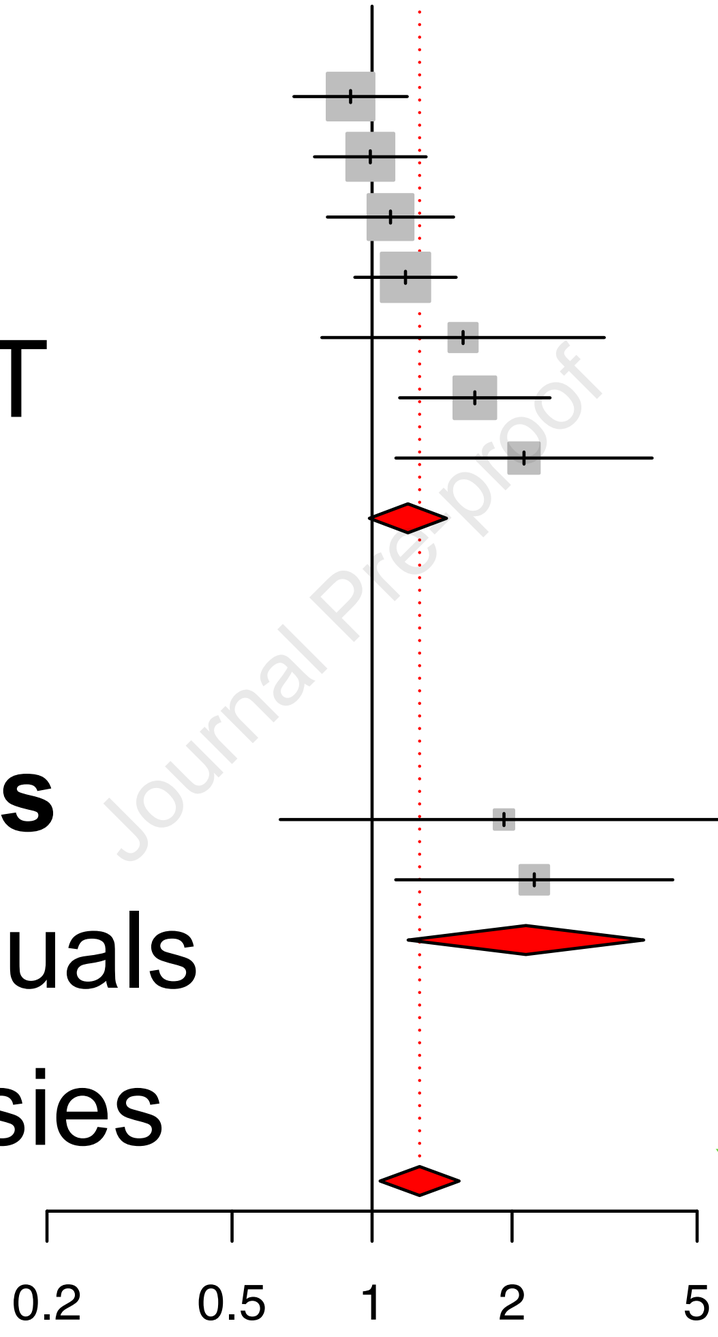
rs641738 C>T

in **NAFLD**

Meta-analysis

>1 million individuals

9,688 liver biopsies



ALT

Liver fat

Fibrosis

Triglycerides

rs641738C>T near MBOAT7 is associated with liver fat, ALT, and fibrosis in NAFLD: a meta-analysis

Kevin Teo[1], Kushala W. M. Abeysekera[2], Leon Adams[3,4], Elmar Aigner[5], Quentin M. Anstee[6,7], Jesus M. Banales[8], Rajarshi Banerjee[9], Priyadarshi Basu[10], Thomas Berg[11], Pallav Bhatnagar[12], Stephan Buch[13], Ali Canbay[14], Sonia Caprio[15], Ankita Chatterjee[10], Yii-Der Ida Chen[16], Abhijit Chowdhury[17], Ann K. Daly[6], Christian Datz[18], Dana de Gracia Hahn[1], Johanna K. DiStefano[19], Jiawen Dong[1], Amedine Duret[1], EU-PNAFLD Investigators*, Connor Emdin[20], Madison Fairey[1], Glenn S Gerhard[21], GOLD Consortium*, Xiuqing Guo[16], Jochen Hampe[13], Matthew Hickman[2], Lena Heintz[22], Christian Hudert[23], Harriet Hunter[1], Matt Kelly[9], Julia Kozlitina[24], Marcin Krawczyk[22,25], Frank Lammert[22], Claudia Langenberg[26], Joel Lavine[27], Lin Li[28], Hong Kai Lim[1], Rohit Loomba[29], Panu K. Luukkonen[30,31,32], Phillip E. Melton[33,34,35], Trevor A. Mori[3], Nicholette D. Palmer[36], Constantinos A. Parisinos[37], Sreekumar G Pillai[12], Faiza Qayyum[38], Matthias C. Reichert[22], Stefano Romeo[39,40,41], Jerome I. Rotter[16], Yu Ri Im[1], Nicola Santoro[15,42], Clemens Schafmayer[43], Elizabeth K. Speliotes[44,45], Stefan Stender[37], Felix Stickel[46], Christopher D. Still[47], Pavel Strnad[48], Kent D. Taylor[16], Anne Tybjærg-Hansen[37], Giuseppina Rosaria Umano[15,48], Mrudula Utukuri[1], Luca Valenti[50,51], Lynne E. Wagenknecht[52], Nicholas J. Wareham[26], Richard M. Watanabe[53], Julia Wattacheril[54], Hanieh Yaghootkar[55], Hannele Yki-Järvinen[30,31], Kendra A. Young[56], Jake P. Mann[26]

Affiliations

1. Clinical School of Medicine, University of Cambridge, Cambridge, UK
2. MRC Integrative Epidemiology Unit (IEU), University of Bristol, Bristol, United Kingdom
3. Medical School, Faculty of Health and Medical Sciences, University of Western Australia, Perth, Australia
4. Department of Hepatology, Sir Charles Gairdner Hospital, Perth, WA, Australia
5. First Department of Medicine, Paracelsus Medical University Salzburg, Austria
6. Translational & Clinical Research Institute, Faculty of Medical Sciences, Newcastle University, Newcastle upon Tyne, United Kingdom
7. Newcastle NIHR Biomedical Research Centre, Newcastle upon Tyne Hospitals NHS Foundation Trust, Newcastle upon Tyne, United Kingdom
8. Department on Liver and Gastrointestinal Diseases, Biodonostia Health Research Institute, Donostia University Hospital, University of the Basque Country (UPV/EHU), CIBERehd, Ikerbasque, San Sebastian, Spain
9. Perspectum Diagnostics Ltd., Oxford, UK
10. National Institute of Biomedical Genomics, Kalyani, India
11. Division of Hepatolgy, Clinic and Polyclinic for Gastroenterology, Hepatology, Infectious Diseases, and Pneumology, University Clinic Leipzig
12. Eli Lilly and Company, Indianapolis, IN

13. Medical Department 1, University Hospital Dresden, Technische Universität Dresden (TU Dresden), Dresden, Germany
14. Gastroenterology, Hepatology and Infectiology, Otto-von-Guericke University Magdeburg, Magdeburg, Germany
15. Yale University, Department of Pediatrics, New Haven, CT
16. The Institute for Translational Genomics and Population Sciences, Department of Pediatrics, The Lundquist Institute for Biomedical Innovation at Harbor-UCLA Medical Center, Torrance, CA
17. Institute of Post Graduate Medical Education And Research, Kolkata, India
18. Department of Internal Medicine, General Hospital Oberndorf, Teaching Hospital of the Paracelsus Medical University Salzburg, Oberndorf, Austria
19. Diabetes and Fibrotic Disease Unit Translational Genomics Research Institute (TGen)
20. Program in Medical and Population Genetics, Broad Institute of Harvard and MIT, Boston, MA
21. Department of Medical Genetics and Molecular Biochemistry, Lewis Katz School of Medicine at Temple University, Philadelphia, PA
22. Department of Medicine II, Saarland University Medical Center, Saarland University, Homburg, Germany
23. Center for Chronically Sick Children, Charité - Universitätsmedizin Berlin, Berlin, Germany
24. Eugene McDermott Center for Human Growth and Development, University of Texas Southwestern Medical Center, Dallas, TX, USA

25. Laboratory of Metabolic Liver Diseases, Department of General, Transplant and Liver Surgery, Laboratory of Metabolic Liver Diseases, Centre for Preclinical Research, Medical University of Warsaw, Warsaw, Poland
26. MRC Epidemiology Unit, Institute of Metabolic Science, University of Cambridge, UK.
27. Department of Pediatrics, Columbia University, New York, USA
28. BioStat Solutions Inc, Frederick, Maryland, USA
29. NAFLD Research Center, Division of Gastroenterology and Epidemiology, University of California at San Diego, La Jolla, California, USA
30. Minerva Foundation Institute for Medical Research, Helsinki, Finland.
31. Department of Medicine, University of Helsinki and Helsinki University Hospital, Helsinki, Finland
32. Yale University School of Medicine, New Haven, Connecticut, US.
33. School of Global Population Health, Faculty of Health and Medical Sciences, The University of Western Australia, Perth, WA, Australia
34. School of Pharmacy and Biomedical Sciences, Faculty of Health Sciences, Curtin University, Perth, WA, Australia
35. Menzies Institute for Medical Research, College of Health and Medicine, University of Tasmania, Hobart, Australia
36. Department of Biochemistry, Wake Forest School of Medicine, Winston-Salem, North Carolina, US
37. Institute of Health Informatics, Faculty of Population Health Sciences, University College London, London, UK

38. Department of Clinical Biochemistry, Rigshospitalet Copenhagen University Hospital Copenhagen, Denmark
39. Department of Molecular and Clinical Medicine, University of Gothenburg, Sweden
40. Cardiology Department, Sahlgrenska University Hospital, Gothenburg, Sweden
41. Clinical Nutrition Unit, Department of Medical and Surgical Sciences, University Magna Graecia, Catanzaro, Italy
42. Department of Medicine and Health Sciences "V. Tiberio" University of Molise, Campobasso, Italy
43. Department of Visceral and Thoracic Surgery, Kiel University, Kiel, Germany
44. Division of Gastroenterology and Hepatology, Department of Medicine, University of Michigan Health System Ann Arbor MI.
45. Department of Computational Medicine and Bioinformatics, University of Michigan Medical School Ann Arbor MI.
46. Department of Gastroenterology and Hepatology, University Hospital of Zurich, Switzerland
47. Geisinger Obesity Institute, Danville, PA 17822, United States
48. Medical Clinic III, University Hospital RWTH Aachen, Aachen, Germany
49. Department of the Woman, the Child, of General and Specialized Surgery, University of Campania Luigi Vanvitelli, Naples Italy
50. Department of Pathophysiology and Transplantation, Università degli Studi di Milano

51. Translational Medicine, Department of Transfusion Medicine and Hematology, Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico Milano, Milan, Italy
52. Division of Public Health Sciences, Wake Forest School of Medicine, Winston-Salem, North Carolina, US
53. Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, California, USA
54. Department of Medicine, Center for Liver Disease and Transplantation, Columbia University College of Physicians and Surgeons, New York Presbyterian Hospital, New York, New York 10032, USA
55. Genetics of Complex Traits, College of Medicine and Health, University of Exeter, Exeter, UK
56. Department of Epidemiology, Colorado School of Public Health, University of Colorado Denver, Aurora, Colorado, USA

*A list of contributors is included in the Acknowledgements

Correspondence: Dr Jake P. Mann, University of Cambridge, Institute of Metabolic Science, Addenbrooke's Hospital, Cambridge, United Kingdom. jm2032@cam.ac.uk, Tel: +44 1223 336792, Fax: +44 1223 330598

Keywords: *MBOAT7*, fibrosis, NAFLD, triglyceride, diabetes, ALSPAC

Word count: 6497

Number of figures: 5

Number of tables: 1

Financial support: JPM is supported by a Wellcome Trust Fellowship (216329/Z/19/Z), a European Paediatric Research Society award, and a Children's Liver Disease Foundation grant. The EU-PNAFLD is supported by an EASL Registry Grant. NIH grants: R01HD028016 (SC), R01DK111038 (SC), R01DK114504 (NS), DK091601 (JKD), UL1TR001105 (JK). Supported by the Nonalcoholic Steatohepatitis Clinical Research Network (NASH CRN) of the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK) (U01DK061718, U01DK061728, U01DK061731, U01DK061732, U01DK061734, U01DK061737, U01DK061738, U01DK061730, and U01DK061713); the National Center for Advancing Translational Sciences (UL1TR000439, UL1TR000436, UL1TR000006, UL1TR000448, UL1TR000100, UL1TR000004, UL1TR000423, UL1TR000058, and UL1TR001881); and the NIDDK (DK063491 to the Southern California Diabetes Endocrinology Research Center). This study was supported by the German Federal Ministry for Education and Research (BmBF) through the Livers Systems Medicine (LiSyM) project. This work was supported by grants from the Swiss National Funds (SNF no. 310030_169196) and the Swiss Foundation for Alcohol Research (SSA) to FS. This Raine Study was supported by the National Health and Medical Research Council of Australia [grant numbers 403981, 353514 and 572613]. The UK Medical Research Council and Wellcome (Grant ref: 102215/2/13/2) and the University of Bristol provide core support for ALSPAC. ALSPAC GWAS data was generated by

Sample Logistics and Genotyping Facilities at Wellcome Sanger Institute and LabCorp (Laboratory Corporation of America) using support from 23andMe. A comprehensive list of grants funding is available on the ALSPAC website (http://www.bristol.ac.uk/alspac/external/documents/grant-acknowledgements.pdf); This research was specifically funded by grants from MRC and Alcohol Research UK (MR/L022206/1) and NIH (5R01AA018333-05) to KWMA & MH. LV was supported by MyFirst Grant AIRC n.16888, Ricerca Finalizzata Ministero della Salute RF-2016-02364358, Ricerca corrente Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico. German Federal Ministry of Education and Research (BMBF LiSyM 031L0051 to F.L.) PL is supported by grants from the Sigrid Jusélius Foundation and the Novo Nordisk Foundation. The Fenland study was funded by grants to the MRC Epidemiology Unit (MC UU12015/1, MC UU 12015/5). RB and MK are employees of and shareholders in Perspectum Diagnostics Ltd. CAP is funded by a Wellcome Trust Clinical PhD Programme (206274/Z/17/Z). JMB is supported by the Spanish Carlos III Health Institute (ISCIII; PI15/01132, PI18/01075, CIBERehd and Miguel Servet Program CON14/00129) co-financed by "Fondo Europeo de Desarrollo Regional" (FEDER), and La Caixa Scientific Foundation (HR17-00601). Research from the GUARDIAN Study was supported DK085175 and DK118062 from the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK). HY is funded by a Diabetes UK RD Lawrence fellowship (17/0005594). QMA and AD supported by the EPoS (Elucidating Pathways of Steatohepatitis) consortium funded by the Horizon 2020 Framework Program of the European Union under Grant Agreement 634413. QMA, AD, LV and AG are members of the LITMUS (Liver

Investigation: Testing Biomarker Utility in Steatohepatitis) consortium funded by the European Union Innovative Medicines Initiative 2 (IMI2) Joint Undertaking under grant agreement 777377. QMA is a Newcastle NIHR Biomedical Research Center Investigator.

Authors' contributions

Study concept and design: JPM; acquisition of data: all; analysis and interpretation of data: KT, JK, SS, LV, JPM; drafting of the manuscript: KT, JPM; critical revision of the manuscript for important intellectual content: all; statistical analysis: KT, JK, SS, LV, JPM; obtained funding: JPM, SC, NS, JKD, JK, PL, JMB, CAP, HY, KWMA, LA, QMA, ADK, TB, GSG, CH, JH, JL, PEM, TAM, NDP, SR, JIR, EKS, SS, ATH, LEW, LV, HYJ, KAY; study supervision: JPM.

Data availability

Raw data used in analyses is available in Supplementary Tables. Code used in analyses is available in the Supplement. Further details are available from the corresponding author upon request.

ABSTRACT

Background & Aims

A common genetic variant near *MBOAT7* (rs641738C>T) has been previously associated with hepatic fat and advanced histology in non-alcoholic fatty liver disease (NAFLD), however, these findings have not been consistently

replicated in the literature. We aimed to establish whether rs641738C>T is a risk factor across the spectrum of NAFLD and characterize its role in the regulation of related metabolic phenotypes through meta-analysis.

Methods

We performed meta-analysis of studies with data on the association between rs641738C>T genotype and: liver fat, NAFLD histology, and serum ALT, lipids, or insulin. These included directly genotyped studies and population-level data from genome-wide association studies (GWAS). We performed random effects meta-analysis using recessive, additive, and dominant genetic models.

Results

Data from 1,066,175 participants (9,688 with liver biopsies) across 42 studies were included in the meta-analysis. rs641738C>T was associated with higher liver fat on CT/MRI (+0.03 standard deviations [95% CI: 0.02 - 0.05], $p_z=4.8x10^{-5}$) and diagnosis of NAFLD (OR 1.17 [95% CI 1.05 - 1.3], $p_z=0.003$) in Caucasian adults. The variant was also positively associated with presence of advanced fibrosis (OR 1.22 [95% CI: 1.03 - 1.45], $p_z=0.021$) in Caucasian adults using a recessive model of inheritance (CC+CT vs. TT). Meta-analysis of data from previous GWAS found the variant to be associated with higher ALT ($p_z=0.002$) and lower serum triglycerides ($p_z=1.5x10^{-4}$). rs641738C>T was not associated with fasting insulin and no effect was observed in children with NAFLD.

Conclusion

Our study validates rs641738C>T near *MBOAT7* as a risk factor for the presence and severity of NAFLD in individuals of European descent.

LAY SUMMARY

Fatty liver disease is a common condition where fat builds up in the liver, which can cause liver inflammation and scarring (including 'cirrhosis'). It is closely linked to obesity and diabetes, but some genes are also thought to be important. We did this study to see whether one specific change ('variant') in one gene (*'MBOAT7'*) was linked to fatty liver disease. We took data from over 40 published studies and found that this variant near *MBOAT7* is linked to more severe fatty liver disease. This means that drugs designed to work on *MBOAT7* may be useful for treating fatty liver disease.

Conflict of interest: Connor Emdin reports personal fees from Navitor Pharma and Novartis.

HIGHLIGHTS

- Meta-analysis of 42 studies (>1 million participants) for the role of rs641738C>T near *MBOAT7* on NAFLD
- rs641738C>T positively associated with liver fat, ALT, fibrosis, and HCC
- rs641738C>T negatively associated with serum triglycerides
- Consistent associations found in studies of Caucasian populations only

INTRODUCTION

Since the first genome-wide association study (GWAS) of liver fat[1], more than 20 genetic single nucleotide variants (SNVs) have been associated with non-alcoholic fatty liver disease (NAFLD)[2]. These studies have deepened our understanding of the condition, its heritability, and its relationship with cardio-metabolic disease.

Rs641738C>T near *MBOAT7* (membrane bound O-acyltransferase domain containing 7) was initially identified as a genome-wide significant risk variant for alcohol-related cirrhosis (odds ratio=1.35, $p=1.03 \times 10^{-9}$) [3], though not replicated in a more recent analysis[4]. It has since been implicated in the pathogenesis of NAFLD[5], hepatocellular carcinoma[6], as well as in fibrosis development in chronic hepatitis B/C[7,8], and primary sclerosing cholangitis[9]. However, unlike variants in *PNPLA3*, *TM6SF2*, and *HSD17B13*, it was not identified at genome-wide significance for liver fat or ALT[1,10,11].

Rs641738 is located a few hundred base pairs downstream of the 3' untranslated region of *MBOAT7*, which belongs to a family of genes that code for specific acyl donors and acceptors[12]. *MBOAT7* encodes lysophosphatidylinositol acyltransferase 1 (LPIAT1), which contributes to the regulation of free arachidonic acid in cells[13,14]. Rs641738C>T is associated with lower hepatic expression of *MBOAT7* at both the mRNA[15] and protein levels[5]. Given its role in inflammatory lipid pathways, most mechanistic work relating to rs641738 has focused on *MBOAT7[16]*.

In NAFLD, the rs641738C>T variant was first demonstrated to be associated with increased hepatic fat content and severity of fibrosis in individuals of European descent[5]. Proton magnetic resonance spectroscopy data from 2,736 individuals showed a modest increase in hepatic fat in those with TT-genotype (4.1%) compared to those with CT- (3.6%) or CC-genotype (3.5%, p=0.005). Follow-up studies of European subjects corroborated the initial findings, and suggested a role in development of hepatocellular carcinoma[17,18]. However, these results were not replicated in adults of other ancestries[5,19–21] or in children[22].

In addition, bi-allelic loss of function mutations in *MBOAT7* cause autosomal recessive mental retardation 57 (OMIM #617188) and no liver phenotype has been reported in these patients to date[14,23]. However, rare likely pathogenic (coding) variants in *MBOAT7* are associated with HCC in NAFLD[24].

In summary, the association between rs641738C>T and hepatic fat content, as well as its effects on severity of NAFLD, remain unclear. Moreover, the broader metabolic effects of this SNV, including its association with markers of insulin resistance and dyslipidaemia have not been assessed. Understanding the broader metabolic effects of rs641738C>T is important if *MBOAT7* were to be investigated as a drug target in NAFLD.

Here, we conducted a large meta-analysis to determine if rs641738C>T influences the development or stage of NAFLD and related traits.

METHODS

Data sources and study selection

Two data sources were included in the meta-analysis: (i) studies which looked at the effect of the variant on traits of interest by genotyping the variant; and (ii) look-up from GWAS of traits of interest.

Studies were sourced through: Medline, Embase, HuGe Navigator, Web of Science, bioRxiv, and medRxiv. The search terms used were: "(*MBOAT7* or membrane-bound-o-acyltransferase) or (rs641738 or rs626283) or (*TMC4*)". In addition, HuGe Navigator Phenopedia was searched using terms related to liver disease (Supplementary Methods). There were no restrictions on date or language. The search was completed on 28th July 2020. Reference lists of publications were also reviewed.

A separate search was conducted for all potentially relevant GWAS through: GWAS Catalogue[25], Phenoscanner[26], Type 2 diabetes knowledge portal[27], and Cardiovascular disease knowledge portal[28] (Supplementary Methods).

After removal of duplicates, titles and abstracts were screened for eligibility independently by two authors (investigators), with inclusion/exclusion criteria applied to potentially eligible full texts.

HuGENet guidelines[29] were followed throughout and MOOSE reporting guidelines[30] were used. This study was prospectively registered on PROSPERO Database of Systematic Reviews (CRD42018105507) available from:

http://www.crd.york.ac.uk/PROSPERO/display_record.php?ID=CRD42018105507

Inclusion and exclusion criteria

Studies were included if genotyping of rs641738C>T (or rs626283G>C [R^2>0.98 in European and American populations[31]] / rs2576452C>T [R^2=0.92 in Guzman *et al.* [32]], which are in strong linkage disequilibrium with rs641738C>T) was conducted and data on one of the outcomes of interest were reported. Narrative review articles, *in vitro* studies, and investigations involving animals, fish, and invertebrates were excluded. Studies which investigated liver disease of other aetiologies were also excluded. There was no restriction on ethnicity or ancestry. Types of studies eligible for inclusion were: case-control, cohort, genome-wide association studies, systematic reviews, and meta-analyses. Pre-print and abstract publications were not eligible for inclusion. Several studies reported on the same cohort (or patient sample) in more than one article. In these instances, data only from the larger of the overlapping cohorts were included in analyses. A full list of overlapping cohorts and articles is in Supplementary Table 1.

Data collection

Details on the recruitment of controls and cases were obtained from each study and, where necessary, clarified by discussion with the study's authors. In particular, it was noted when cases and controls were not recruited from the same population or clinics.

Hepatic steatosis or NAFLD (as diagnosis) was evaluated as a dichotomous variable where radiological (liver ultrasound, controlled attenuation parameter [CAP, with cut-off >248dB/m], CT, MRI) or histological assessment were used. Hepatic fat content was collected as a continuous variable from CT, MRS, MRI, PDFF. Non-invasive assessment of hepatic fat content was also assessed using semi-quantitative scoring in the Fenland cohort, as previously described[33], and using CAP.

Individual participant-level histology data were extracted according to the NASH Clinical Research Network scoring system[34] and, where not otherwise diagnosed by a pathologist's assessment, NASH was defined using the Fatty Liver Inhibition of Progression (FLIP) algorithm[35]. The above data were collected for each genotype separately (CC, CT, and TT).

Participant demographics and characteristics meta-data were collected from each study, including: sex, age, ethnicity, presence of type 2 diabetes, body mass index (BMI). Where possible, individual patient-level data was obtained. The authors of 59 studies were contacted for additional data or clarification, of whom 49 replied. Data from 11 potentially relevant studies could not be included, which are listed in the Supplementary Methods.

Additional details regarding cohorts with genome-wide data, the Avon Longitudinal Study of Parents and Children (ALSPAC) [36–38]data extracted from the UK BioBank, quality assessment, and statistical analysis is found in the Supplementary Methods.

RESULTS

Database search identified 1167 articles (Supplementary Fig. 1), of which 44 articles were included: 42 primary studies (Supplementary Tables 2-4), one systematic review and one meta-analysis (Supplementary Table 5).
In total, 1,066,175 individuals (5,711 children) were included in the meta-analysis. Most studies were in adults (32/42, 76%) and in individuals from predominantly Caucasian populations (26/42, 62%). Of the 42 included studies, 14 studies (9,688 participants, hereof 584 children) reported data on liver histology.

Studies were generally of high quality, though in five studies[11,22,39–41] (four in adults and one in children) the control group was recruited from a different population or sample to the cases (Supplementary Table 3).

One previous meta-analysis was included[42], which used data from 5 case-control studies to assess the effect of rs641738C>T on diagnosis of NAFLD. The meta-analysis included 2,560 cases and 8,738 controls and found no evidence of an association between this variant and diagnosis of NAFLD (Supplementary Table 5). One previous systematic review[43] found positive associations between rs641738C>T in adults of Caucasian, Hispanic, and African American descent with limited data in children (Supplementary Table 5).

Liver fat, NAFLD, and severe steatosis in adults

Seven studies (29,679 participants) reported data on hepatic fat as a continuous variable assayed by CT or MRI. On meta-analysis, rs641738C>T was associated with higher liver fat in studies in Caucasian populations using an additive model of inheritance, with a per T-allele change of β 0.034 (95% CI 0.018, 0.051), $p_z=4.8x10^{-5}$) standard deviations in inverse-normalized liver fat (Figure 1), whilst no consistent effect was observed in non-Caucasian populations. A similar trend was observed using a dominant model of inheritance in studies of Caucasian populations: mean difference in hepatic fat +0.18% ((95% CI 0.2, 0.34), $p_z=0.04$, Supplementary Table 6).

Given the difference in sensitivity and specificity of modalities used to assess liver fat, a sub-analysis by modality of imaging was performed. No significant differences were observed between studies using CT, MRI, or MRS for quantification of liver fat (Supplementary Figure 2).

A similar trend was observed using CAP and semi-quantitative ultrasound to assess steatosis severity in 12,224 adults (β 0.02 (95% CI -0.002, 0.04), , $p_z=0.08$, Supplementary Figure 3).

Data from a range of diverse modalities was used to assess the effect of this variant on diagnosis of NAFLD, to reflect real-world diagnostic practice. rs641738C>T was associated with NAFLD as a trait (OR 1.15 (95% CI 1.05, 1.26), $p_z=0.002$) using a recessive model of inheritance (Figure 2) but not using additive or dominant models (Supplementary Table 7). The effect was only observed in studies of Caucasian populations (OR 1.17 (95% CI 1.05,

1.3), p_z=0.003). Sub-group analysis by modality of diagnosis found the 95% confidence intervals for all modalities overlapped, except for MRI-PDFF, which had only one study (Supplementary Figure 4). The association remained after excluding four studies where there was a lack of similarity between cases and controls (OR 1.19 (95% CI 1.07, 1.33), p_z=0.0017) using a recessive model of inheritance.

However, Egger's test suggested evidence of study distribution (publication) bias (p=0.013) and when using the Trim and Fill method to account for this bias, the positive association remained but was attenuated (OR 1.11 (95% CI 1.01, 1.23), p_z=0.037, Supplementary Figure 5).

In patients with NAFLD, data from eight studies (6,206 participants) rs641738C>T was not significantly associated with the presence of severe steatosis (S1-2 vs. S3) on liver biopsy (OR 1.08 (95% CI 0.78, 1.5), P_z=0.64, Table 1 & Supplementary Figure 6).

Histological NASH in adults

Data from nine studies (7,719 participants) found that rs641738C>T was not associated with the presence of NASH on biopsy in adults (OR 1.24 (95% 0.96, 1.36), p_z=0.128 Supplementary Figure 7).

Fibrosis in adults

Liver biopsy data on presence of advanced fibrosis was available from eight studies (7,692 adults). Our primary outcome, presence of advanced fibrosis in

adults (stage F0-2 versus stage F3-4), showed a borderline positive association with rs641738C>T in Caucasian populations (OR 1.22 (95% 1.03, 1.45), p_z=0.021)(Figure 3). In addition, two studies used ICD-codes (International Statistical Classification of Diseases and Related Health Problems) in the UKBB cohort to identify individuals with NAFLD and advanced fibrosis or cirrhosis[44,45]. Both found positive associations below genome-wide significance: for example, using an additive model of inheritance Emdin *et al.* found the association between rs641738C>T and cirrhosis as β 1.22 (SE 0.06, P=0.03), using an additive genetic model.

Data from nine studies (8,389 participants), found that presence of any fibrosis (F0 versus F1-4) was also borderline positively associated with rs641738C>T overall (OR 1.27 (95% 1.04, 1.54), p_z=0.018) as well as in non-Caucasian populations as a sub-group (Supplementary Figure 8).

Development of hepatocellular carcinoma

Four cohorts (2,328 participants, 228 cases of NAFLD-HCC) reported on development of HCC in patients with NAFLD. rs641738C>T was associated with increased odds of HCC in NAFLD only when using a dominant model (CC vs. CT+TT) of inheritance (OR 1.64 (95% CI 1.18, 2.27), p_z=0.003, Figure 4).

Effect on alanine aminotransferase (ALT)

Data from GWAS using log-transformed ALT (609,794 participants) were available for meta-analysis to investigate the role of rs641738C>T on ALT.

The variant showed a positive association with ALT (β 0.004 (95% CI 0.002, 0.007), p_z=0.002), which on sub-analysis was observed in Caucasian populations but not in non-Caucasian populations (Figure 5 & Supplementary Table 7).

Additionally, in the UKBB cohort, rs641738C>T was associated with a small, but statistically significant (P=2.0x10^{-8}) increase in un-transformed ALT: 0.18 IU/L higher ALT per T-allele in this variant (Supplementary Table 8).

In the remaining cohort and case-control studies included in the meta-analysis (15,208 adults), rs641738C>T was not found to be significantly associated with a change in ALT, for example mean difference using a recessive model (CC+CT vs. TT) +0.32 IU/L (95% CI -0.06, 0.7), p_z=0.08, Supplementary Table 9) in Caucasian populations.

Effect on serum lipids and insulin

Data from GWAS using log-transformed serum triglycerides (850,241 participants) found that rs641738C>T was associated with lower triglycerides (β -0.01 (95% CI -0.018, -0.006), p_z=1.5x10^{-4}), which on sub-analysis was observed in Caucasian populations but not in non-Caucasian populations (Supplementary Fig. 9). Similar findings were obtained from meta-analysis of cohort and case control studies, particularly using an additive model (β -0.03 (95% CI -0.05, -0.01), p_z=0.00091.), Supplementary Table 9).

Data from GWAS (852,409 participants) found rs641738C>T to be positively associated with total cholesterol , in Caucasian populations (β 0.007 (95% CI 0.003, 0.01), $p_z=2.1x10^{-4}$), which was not observed in non-Caucasian populations (Supplementary Fig. 10). A borderline positive association was also observed between rs641738C>T and high-density lipoprotein (HDL) cholesterol (β 0.009 (95% CI 0.001, 0.02), $p_z=0.02$), Supplementary Table 7). There was no effect on fasting insulin levels found in population-level GWAS (β 0.009 (95% CI -0.03, 0.04), $p_z=0.64$), Supplementary Table 7). However, a negative association was observed using data from cohort and case-control studies with a dominant genetic model (mean difference -1.4 pmol/L (95% CI -2.1, -0.65), $p_z=0.004$), Supplementary Table 9).

Effect of rs641738C>T on paediatric NAFLD

Data from ten studies (5,711 children) was used in the meta-analysis. rs641738C>T was not significantly associated with the diagnosis of NAFLD, liver fat content, stage of liver histology, or serum biochemistry (Supplementary table 10).

Meta-regression shows interaction between rs641738C>T and type 2 diabetes

Finally, we aimed to determine whether baseline participant characteristics influenced the association of rs641738C>T on histological outcomes using meta-regression. There was a negative association with presence of type 2 diabetes and effect size for NASH vs. NAFL (β -1.8 (standard error 0.65), p=0.006, Supplementary Figure 11A). A similar negative trend with type 2

diabetes was observed for severe steatosis (S1-2 vs. S3, β -2.6 (standard error 1.5), p=0.08) and presence of fibrosis (F0 vs. F1-4, β -1.5 (standard error 0.8), p=0.06, Supplementary Table 11). In addition, effect size for any fibrosis was greater in cohorts with an older mean age (β 0.05 (standard error 0.02), p=0.014, Supplementary Figure 11D).

DISCUSSION

Identification of genetic variants associated with NAFLD has the potential to inform pre-clinical research and our understanding of hepatic metabolism. In this meta-analysis we have validated rs641738C>T near *MBOAT7* as a risk factor for the full spectrum of NAFLD in Caucasian adults.

A two-stage GWAS initially identified rs641738C>T as a genome-wide significant locus for alcohol-related cirrhosis[3]. *MBOAT7* was a potentially interesting target as an enzyme involved in (phospho)lipid metabolism, conceptually similar to other SNVs at GWAS-significance in alcoholic and non-alcoholic liver disease, namely *TM6SF2* and *PNPLA3*. Later studies found the variant to influence the full spectrum of fatty liver disease, from steatosis to NASH, to fibrosis, cirrhosis and HCC[5,17]. However, these associations have not been consistently replicated in the literature[19]. We conducted a meta-analysis to firmly establish the association of rs641738C>T with the presence and severity of NAFLD, and associated metabolic traits.

Main findings

We found that rs641738C>T was associated with higher liver fat content, higher ALT, and with higher odds of NAFLD diagnosis, fibrosis, and HCC, particularly in Caucasian adults and in the homozygous 'TT' genotype. The effects sizes of rs641738C>T reported here are small compared to those of *PNPLA3* p.I148M and *TM6SF2* p.E167K, the two strongest steatogenic variants[46]. Also, the magnitude of change in alanine aminotransferase is small relative to that associated with variants in *PNPLA3*, *HSD17B13*,

MTARC1, and *TM6SF2*. This may account for the absence of this variant (or others near *MBOAT7*) from GWAS for NAFLD in the general population[1,10,11,45,47] the effect size (and associated p-value) was too small to be identified as significant genome-wide. The marginal positive effect on hepatic triglyceride content may suggest this variant acts through alterations in the composition of hepatic lipid, as well as quantity[17]. This is consistent with pre-clinical data on lipotoxicity, where the composition of hepatic fats influence development of NASH. On the other hand, a recent Mendelian randomization study using these variables as instruments to assess causality of fatty liver in determining fibrosis has shown the effect of steatosis highly correlates with fibrosis in all the genetic variables indicating that quantity of lipid rather than quality may be more important[48]. Functional studies are needed to understand the relationship between quality/quantity of fat and hepato-toxic/-protective mechanism in causing progression of disease.

The function of this variant is still relatively poorly understood and there is conflicting evidence as to whether rs641738C>T is associated with changes in hepatic expression of *MBOAT7*. Results from the GTEx Consortium show a strong negative association with T-allele[15], which is supported by data from Schadt *et al.[49]*. *MBOAT7* protein expression correlated with mRNA in liver biopsies from Mancina *et al.[5]* but this finding was not replicated by Sookoian *et al.[19]*. *MBOAT7* encodes LPIAT1, a 6 transmembrane domain protein involved in acyl-chain remodelling of membranes that influence intracellular membrane composition and circulating phosphatidylinositols[50]. Further, recent metabolite profiling data implicates *MBOAT7* as the causal gene for

this SNV[51]. Moreover, TMC4 was found with a low expression in the liver[5] that is consistent with no mechanistic data supporting its role in NAFLD.

The hypothesis that *MBOAT7* is the causal gene underlying the association with liver disease at the locus is supported by the observation that mice deficient for *MBOAT7* have altered hepatic concentrations of polyunsaturated phosphatidylinositol[50]. Similarly, metabolite data from humans is strongly suggestive that rs641738C>T reduces MBOAT7 function[52]. In addition, two independent groups have found that loss of *MBOAT7* (but not *TMC4)* increases the severity of NAFLD in mice fed a high-fat diet[53,54].

These analyses suggest that rs641738C>T impacts the severity of NAFLD through a recessive model of inheritance, though some analyses (e.g. liver fat and ALT) were suggestive of a role using an additive genetic model. Other genetic variants are known to impact on all-cause mortality in a recessive manner, notably variants that perturb *HFE[44]*. Further mechanistic work is required to understand the extent to which the haplo-insufficient state affects hepatocyte function.

We found no evidence of an effect of rs641738C>T on insulin resistance: the key driver of hepatic steatosis, as determined by unaltered fasting insulin concentrations. GWAS meta-analyses of type 2 diabetes have implicated p.I148M in *PNPLA3* and p.E167K in *TM6SF2* as significant risk loci (albeit with very modest effect size as compared to their effects on liver disease)[55] and Mendelian randomization studies indicate a causal role in determining

insulin resistance mediated by the degree of liver damage[48,56]. Similarly, these two variants are associated with reduced risk of coronary artery disease; whilst our analysis did find lower serum triglycerides to be associated with this variant it has not been associated with lower rates of cardiovascular disease[57]. However, we did observe a negative association between effect size and prevalence of diabetes on meta-regression, potentially suggesting that this variant has the greatest effect in less insulin-resistant individuals.

A strength of this meta-analysis is the large number of individuals with liver biopsy-derived phenotypic data as well as use of population-based GWAS data. The larger number of included studies and participants is likely to account for the different conclusions reached in this study compared to the previous meta-analysis by Xia *et al.[42]*.

Limitations and quality of evidence

An important practical consideration is the population frequency of this variant in different ethnicities. The mean allelic frequency of the effect (T) allele is highly variable: from 0.24 in East Asians compared to 0.53 in those of South Asian ancestry[58]. Moreover, the majority of studies included in this meta-analysis used self-reported ethnicity, rather than genetic ancestry.

Though this analysis did include data from individuals of multiple ethnicities (and genetic ancestries) we only found evidence of an effect of this variant in Caucasian individuals. This is consistent with the initial discovery and it is likely that rs641738C>T is a proxy for the true causal variant. However, due to

differences in patterns of linkage disequilibrium, we cannot exclude the possibility that a different nearby locus is associated with liver-related phenotypes in individuals of other genetic ancestries.

A limitation of using meta-analysis for a single variant is the lack of adjustment for population stratification. When further genome-wide data are available, a formal GWAS meta-analysis may be able to address this.

We found significant differences between adult and paediatric histological analyses. Whilst there were fewer clinical events (e.g. with advanced fibrosis) in children, the analyses did not show a trend congruous with those in adults. Paediatric NAFLD has a different histological phenotype to that of adults (with prominent periportal inflammation) and it is therefore plausible that this is a true lack of association in children with NAFLD.

Data from multiple diagnostic or imaging modalities were combined in several analyses. Though we observed minimal heterogeneity between modalities, these techniques have differing accuracy for diagnosis of steatosis, which has the potential to affect results. The sub-group analysis of hepatic fat by modality suggested a marginally greater effect size in studies using MRS, which is regarded as a highly sensitive technique. There is potential that through inclusion of other modalities (e.g. CT) we have underestimated the effect size associated with this variant.

The magnitude of effect observed across all associations is small in comparison to other well-established variants. The clinical relevance of rs738409C>G in *PNPLA3* has been validated with hard end-points[59] but large cohorts will be required to prospectively demonstrate the clinical risk associated with this variant near *MBOAT7*.

Though there was minimal heterogeneity across included studies, there was evidence of publication bias but the effect on diagnosis of NAFLD appeared to persist after attempting to account for this. Also of note, the numbers of individuals with NAFLD and HCC were comparatively low, also limiting the power to assess for an association of this variant with non-cirrhotic HCC, as has been previously reported[6]. The HCC analysis was also unique in only demonstrating an effect in the dominant, rather than recessive, model of inheritance. Further work in this area may improve the accuracy of effect estimates.

Conclusions

rs641738C>T near *MBOAT7* is positively associated with liver fat, ALT, and histological severity in Caucasian adults with NAFLD, but negatively associated with serum triglycerides and with relatively small effect sizes throughout. These data validate this locus as significant in the pathogenesis of NAFLD.

ABBREVIATIONS

ALSPAC, Avon Longitudinal Study of Parents and Children; ALT, alanine aminotransferase; BMI, body mass index; CAP, controlled attenuation parameter; CT, computed tomography; GWAS, genome-wide association study; *HFE*, homeostatic iron regulator protein; HOMA-IR, homeostatic model assessment of insulin resistance; HSD17B13, 17β-Hydroxysteroid dehydrogenase type 13; *MBOAT7*, membrane bound O-acyltransferase domain containing 7; *MTARC1*, Mitochondrial amidoxime reducing component 1; OR, odds ratio; MRI-PDFF, magnetic resonance imaging-proton density fat fraction; MRS, magnetic resonance spectroscopy; *PNPLA3*, patatin-like phospholipase domain containing protein 3; SNV, single nucleotide variant; *TM6SF2*, transmembrane 6 superfamily member 2; *TMC4*, transmembrane channel-like 4; UKBB, UK BioBank.

ACKNOWLEDGEMENTS

The authors are grateful to the Raine Study participants and their families, and to the Raine Study research staff for cohort coordination and data collection. The authors gratefully acknowledge the following institutes for providing funding for Core Management of the Raine Study: The University of Western Australia (UWA), Curtin University, the Raine Medical Research Foundation, the UWA Faculty of Medicine, Dentistry and Health Sciences, the Telethon Kids Institute, the Women and Infants Research Foundation (King Edward Memorial Hospital) and Edith Cowan University).

We are extremely grateful to all the families who took part in the ALSPAC study, the midwives for their help in recruiting them, and the whole ALSPAC team, which includes interviewers, computer and laboratory technicians,

clerical workers, research scientists, volunteers, managers, receptionists and nurses.

This study has been conducted using data from the Fenland study. The authors gratefully acknowledge the help of the MRC Epidemiology Unit Support Teams, including Field, Laboratory and Data Management Teams. The authors are grateful to the members of the EU-PNAFLD Registry, including Anita Vreugdenhil, Anna Alisi, Piotr Socha, Wojciech Jańczyk, Ulrich Baumann, Sanjay Rajwal, Indra van Mourik, Florence Lacaille, Myriam Dabbas, Deirdre A. Kelly, and the late Valerio Nobili. We would also like to thank Naga Chalasani for his helpful comments. This research has made use of the UK Biobank resource under project number 9914.

REFERENCES

Author names in bold designate shared co-first authorship.

[1] **Speliotes EK, Yerges-armstrong LM, Wu J, Hernaez R, Kim LJ**, Palmer CD, et al. Genome-Wide Association Analysis Identifies Variants Associated with Nonalcoholic Fatty Liver Disease That Have Distinct Effects on Metabolic Traits. PLoS Genet 2011;7:e1001324.

[2] Romeo S, Sanyal A, Valenti L. Leveraging Human Genetics to Identify Potential New Treatments for Fatty Liver Disease. Cell Metab 2020;31:35–45.

[3] Buch S, Stickel F, Trépo E, Way M, Herrmann A, Nischalke HD, et al. A genome-wide association study confirms PNPLA3 and identifies TM6SF2 and MBOAT7 as risk loci for alcohol-related cirrhosis. Nat Genet 2015;47:1443–8.

[4] **Innes H, Buch S**, Hutchinson S, Guha IN, Morling JR, Barnes E, et al. Genome-wide Association Study for Alcohol-related Cirrhosis Identifies Risk Loci in MARC1 and HNRNPUL1. Gastroenterology 2020. https://doi.org/10.1053/j.gastro.2020.06.014.

[5] **Mancina RM, Dongiovanni P**, Petta S, Pingitore P, Meroni M, Rametta R, et al. The MBOAT7-TMC4 Variant rs641738 Increases Risk of Nonalcoholic Fatty Liver Disease in Individuals of European Descent. Gastroenterology 2016;150:1219–30e6.

[6] **Donati B, Dongiovanni P**, Romeo S, Meroni M, McCain M, Miele L, et al. MBOAT7 rs641738 variant and hepatocellular carcinoma in non-cirrhotic individuals. Sci Rep 2017;7:4492.

[7] Thabet K, Asimakopoulos A, Shojaei M, Romero-Gomez M, Mangia A,

Irving WL, et al. MBOAT7 rs641738 increases risk of liver inflammation and transition to fibrosis in chronic hepatitis C. Nat Commun 2016;7:12757.

[8] Thabet K, Chan HLY, Petta S, Mangia A, Berg T, Boonstra A, et al. The membrane-bound O-acyltransferase domain-containing 7 variant rs641738 increases inflammation and fibrosis in chronic hepatitis B. Hepatology 2017;65:1840–50.

[9] Freund C, Wahlers A, Begli NH, Leopold Y, Klöters-Plachky P, Mehrabi A, et al. The MBOAT7 rs641738 variant is associated with an improved outcome in primary sclerosing cholangitis. Clin Res Hepatol Gastroenterol 2020. https://doi.org/10.1016/j.clinre.2019.12.006.

[10] Abul-Husn NS, Cheng X, Li AH, Xin Y, Schurmann C, Stevis P, et al. A Protein-Truncating *HSD17B13* Variant and Protection from Chronic Liver Disease. N Engl J Med 2018;378:1096–106.

[11] Anstee QM, Darlay R, Cockell S, Meroni M, Govaere O, Tiniakos D, et al. Genome-wide association study of non-alcoholic fatty liver and steatohepatitis in a histologically-characterised cohort. J Hepatol 2020. https://doi.org/10.1016/j.jhep.2020.04.003.

[12] Matsuda S, Inoue T, Lee HC, Kono N, Tanaka F, Gengyo-Ando K, et al. Member of the membrane-bound O-acyltransferase (MBOAT) family encodes a lysophospholipid acyltransferase with broad substrate specificity. Genes Cells 2008;13:879–88.

[13] **Gijón MA, Riekhof WR**, Zarini S, Murphy RC, Voelker DR. Lysophospholipid acyltransferases and arachidonate recycling in human neutrophils. J Biol Chem 2008;283:30235–45.

[14] Johansen A, Rosti RO, Musaev D, Sticca E, Harripaul R, Zaki M, et al. Mutations in MBOAT7, Encoding Lysophosphatidylinositol Acyltransferase I, Lead to Intellectual Disability Accompanied by Epilepsy and Autistic Features. Am J Hum Genet 2016;99:912–6.

[15] The GTEx Consortium. The Genotype-Tissue Expression (GTEx) project. Nat Genet 2013;45:580–5.

[16] Meroni M, Dongiovanni P, Longo M, Carli F, Baselli G, Rametta R, et al. Mboat7 down-regulation by hyper-insulinemia induces fat accumulation in hepatocytes. EBioMedicine 2020;52:102658.

[17] Luukkonen PK, Zhou Y, Hyötyläinen T, Leivonen M, Arola J, Orho-Melander M, et al. The MBOAT7 variant rs641738 alters hepatic phosphatidylinositols and increases severity of non-alcoholic fatty liver disease in humans. J Hepatol 2016;65:1263–5.

[18] Krawczyk M, Rau M, Schattenberg JM, Bantel H, Pathil A, Demir M, et al. Combined effects of the *PNPLA3* rs738409, *TM6SF2* rs58542926, and *MBOAT7* rs641738 variants on NAFLD severity: a multicenter biopsy-based study. J Lipid Res 2017;58:247–55.

[19] Sookoian S, Flichman D, Garaycoechea ME, Gazzi C, Martino JS, Castaño GO, et al. Lack of evidence supporting a role of TMC4-rs641738 missense variant - MBOAT7- intergenic downstream variant - In the Susceptibility to Nonalcoholic Fatty Liver Disease. Sci Rep 2018;8:5097.

[20] **Koo BK, Joo SK**, Kim D, Bae JM, Park JH, Kim JH, et al. Additive effects of PNPLA3 and TM6SF2 on the histological severity of non-alcoholic fatty liver disease. J Gastroenterol Hepatol 2018;33:1277–85.

[21] Umano GR, Caprio S, Di Sessa A, Chalasani N, Dykas DJ, Pierpont B, et

al. The rs626283 variant in the MBOAT7 gene is associated with insulin resistance and fatty liver in Caucasian obese youth. Am J Gastroenterol 2018;113:376–83.

[22] Hudert CA, Selinski S, Rudolph B, Bläker H, Loddenkemper C, Thielhorn R, et al. Genetic determinants of steatosis and fibrosis progression in paediatric non-alcoholic fatty liver disease. Liver Int 2019;39:540–56.

[23] Yalnızoğlu D, Özgül RK, Oğuz KK, Özer B, Yücel-Yılmaz D, Gürbüz B, et al. Expanding the phenotype of phospholipid remodelling disease due to MBOAT7 gene defect. J Inherit Metab Dis 2019;42:381–8.

[24] **Pelusi S, Baselli G**, Pietrelli A, Dongiovanni P, Donati B, McCain MV, et al. Rare Pathogenic Variants Predispose to Hepatocellular Carcinoma in Nonalcoholic Fatty Liver Disease. Sci Rep 2019;9:3682.

[25] MacArthur J, Bowler E, Cerezo M, Gil L, Hall P, Hastings E, et al. The new NHGRI-EBI Catalog of published genome-wide association studies (GWAS Catalog). Nucleic Acids Res 2017;45:D896–901.

[26] Kamat MA, Blackshaw JA, Young R, Surendran P, Burgess S, Danesh J, et al. PhenoScanner V2: an expanded tool for searching human genotype-phenotype associations. Bioinformatics 2019;35:4851–3.

[27] Type 2 diabetes knowledge portal n.d. http://www.type2diabetesgenetics.org/ (accessed July 28, 2020).

[28] Cardiovascular Disease Knowledge Portal n.d. http://www.broadcvdi.org/ (accessed July 28, 2020).

[29] Little J, Higgins J (editors). The HuGENet™ HuGE Review Handbook, version 1.0. Centers for Disease Control and Prevention; 2006.

[30] Stroup DF, Berlin JA, Morton SC, Olkin I, Williamson GD, Rennie D, et al.

Meta-analysis of Observational Studies. JAMA 2000.

[31] **Arnold M, Raffler J**, Pfeufer A, Suhre K, Kastenmüller G. SNiPA: an interactive, genetic variant-centered annotation browser. Bioinformatics 2015;31:1334–6.

[32] Guzman CB, Duvvuru S, Akkari A, Bhatnagar P, Battioui C, Foster W, et al. Coding variants in PNPLA3 and TM6SF2 are risk factors for hepatic steatosis and elevated serum alanine aminotransferases caused by a glucagon receptor antagonist. Hepatology Communications 2018;2:561–70.

[33] De Lucia Rolfe E, Brage S, Sleigh A, Finucane F, Griffin SJ, Wareham NJ, et al. Validity of ultrasonography to assess hepatic steatosis compared to magnetic resonance spectroscopy as a criterion method in older adults. PLoS One 2018;13:87–99.

[34] Kleiner DE, Brunt EM, Van Natta M, Behling C, Contos MJ, Cummings OW, et al. Design and validation of a histological scoring system for nonalcoholic fatty liver disease. Hepatology 2005;41:1313–21.

[35] Bedossa P, Burt AA, Gouw AHA, Lackner C, Schirmacher P, Terracciano L, et al. Utility and appropriateness of the fatty liver inhibition of progression (FLIP) algorithm and steatosis, activity, and fibrosis (SAF) score in the evaluation of biopsies of nonalcoholic fatty liver disease. Hepatology 2014;60:565–75.

[36] Boyd A, Golding J, Macleod J, Lawlor DA, Fraser A, Henderson J, et al. Cohort Profile: the “children of the 90s”--the index offspring of the Avon Longitudinal Study of Parents and Children. Int J Epidemiol 2013;42:111–27.

[37] Northstone K, Lewcock M, Groom A, Boyd A, Macleod J, Timpson N, et al. The Avon Longitudinal Study of Parents and Children (ALSPAC): an update on the enrolled sample of index children in 2019. Wellcome Open Res 2019;4:51.

[38] Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009;42:377–81.

[39] Di Costanzo A, Belardinilli F, Bailetti D, Sponziello M, D'Erasmo L, Polimeni L, et al. Evaluation of Polygenic Determinants of Non-Alcoholic Fatty Liver Disease (NAFLD) By a Candidate Genes Resequencing Strategy. Sci Rep 2018;8:1–10.

[40] Kawaguchi T, Shima T, Mizuno M, Mitsumoto Y, Umemura A, Kanbara Y, et al. Risk estimation model for nonalcoholic fatty liver disease in the Japanese using multiple genetic markers. PLoS One 2018;13:1–16.

[41] Koo BK, Joo SK, Kim D, Lee S, Bae JM, Park JH, et al. Development and Validation of a Scoring System, Based on Genetic and Clinical Factors, to Determine Risk of Steatohepatitis in Asian Patients with Nonalcoholic Fatty Liver Disease. Clin Gastroenterol Hepatol 2020. https://doi.org/10.1016/j.cgh.2020.02.011.

[42] Xia Y, Huang CX, Li GY, Chen KH, Han L, Tang L, et al. Meta-analysis of the association between MBOAT7 rs641738, TM6SF2 rs58542926 and nonalcoholic fatty liver disease susceptibility. Clin Res Hepatol Gastroenterol 2019:1–9.

[43] Ismaiel A, Dumitrascu DL. Genetic predisposition in metabolic-

dysfunction-associated fatty liver disease and cardiovascular outcomes-Systematic review. Eur J Clin Invest 2020:e13331.

[44] Emdin CA, Haas ME, Khera AV, Aragam K, Chaffin M, Klarin D, et al. A missense variant in Mitochondrial Amidoxime Reducing Component 1 gene and protection against liver disease. PLoS Genet 2020;16:e1008629.

[45] Chen VL, Chen Y, Du X, Handelman SK, Speliotes EK. Genetic variants that associate with cirrhosis have pleiotropic effects on human traits. Liver Int 2020;40:405–15.

[46] **Sookoian S, Pirola CJ, Valenti L, Davidson NO**. Genetic Pathways in Nonalcoholic Fatty Liver Disease: Insights From Systems Biology. Hepatology 2020;72:330–46.

[47] Chambers JC, Zhang W, Sehmi J, Li X, Wass MN, Van Der Harst P, et al. Genome-wide association study identifies loci influencing concentrations of liver enzymes in plasma. Nat Genet 2011;43:1131–8.

[48] **Dongiovanni P, Stender S, Pietrelli A, Mancina RM**, Cespiati A, Petta S, et al. Causal relationship of hepatic fat with liver damage and insulin resistance in nonalcoholic fatty liver. J Intern Med 2018;283:356–70.

[49] **Schadt EE, Molony C, Chudin E**, Hao K, Yang X, Lum PY, et al. Mapping the genetic architecture of gene expression in human liver. PLoS Biol 2008;6:1020–32.

[50] Lee H-C, Inoue T, Sasaki J, Kubo T, Matsuda S, Nakasaki Y, et al. LPIAT1 regulates arachidonic acid content in phosphatidylinositol and is required for cortical lamination in mice. Mol Biol Cell 2012;23:4689–700.

[51] **Mann JP, Pietzner M**, Wittemans LB, De Lucia Rolfe E, Nicola D,

Imamura F, et al. Insights into genetic variants associated with NASH-fibrosis from metabolite profiling. Hum Mol Genet 2020:doi: 10.1093/hmg/ddaa162.

[52] **Shin S-Y, Fauman EB, Petersen A-K, Krumsiek J**, Santos R, Huang J, et al. An atlas of genetic influences on human blood metabolites. Nat Genet 2014;46:543–50.

[53] Helsley RN, Varadharajan V, Brown AL, Gromovsky AD, Schugar RC, Ramachandiran I, et al. Obesity-linked suppression of membrane-bound O-acyltransferase 7 (MBOAT7) drives non-alcoholic fatty liver disease. Elife 2019;8. https://doi.org/10.7554/eLife.49882.

[54] **Tanaka Y, Shimanaka Y, Caddeo A**, Kubo T, Mao Y, Kubota T, et al. LPIAT1/MBOAT7 depletion increases triglyceride synthesis fueled by high phosphatidylinositol turnover. Gut 2020. https://doi.org/10.1136/gutjnl-2020-320646.

[55] Mahajan A, Taliun D, Thurner M, Robertson NR, Torres JM, Rayner NW, et al. Fine-mapping type 2 diabetes loci to single-variant resolution using high-density imputation and islet-specific epigenome maps. Nat Genet 2018;50:1505–13.

[56] **Parisinos CA, Wilman HR**, Thomas EL, Kelly M, Nicholls RC, McGonigle J, et al. Genome-wide and Mendelian randomisation studies of liver MRI yield insights into the pathogenesis of steatohepatitis. J Hepatol 2020;73:241–51.

[57] Brouwers MCGJ, Simons N, Stehouwer CDA, Koek GH, Schaper NC, Isaacs A. Relationship Between Nonalcoholic Fatty Liver Disease Susceptibility Genes and Coronary Artery Disease. Hepatol Commun

2019;3:587–96.

[58] Karczewski KJ, Francioli LC, Tiao G, Cummings BB, Alföldi J, Wang Q, et al. The mutational constraint spectrum quantified from variation in 141,456 humans. Nature 2020;581:434–43.

[59] **Grimaudo S, Pipitone RM**, Pennisi G, Celsa C, Cammà C, Di Marco V, et al. Association Between PNPLA3 rs738409 C>G Variant and Liver-Related Outcomes in Patients With Nonalcoholic Fatty Liver Disease. Clin Gastroenterol Hepatol 2020;18:935–44.e3.

Tables

Outcome	Genetic model	Sub-analysis	No. of studies	Heterogeneity		Effect summary	
				I^2	p_Q	OR [95% CI]	p_Z
NAFLD diagnosis (control vs NAFLD)	Recessive	Overall	17	0.25	0.17	1.15 (1.05, 1.26)	0.0018
NAFLD diagnosis (control vs NAFLD)	Recessive	Non-Caucasian	5	0	0.46	1.1 (.9, 1.34)	0.343
NAFLD diagnosis (control vs NAFLD)	Recessive	Caucasian	12	0.38	0.09	1.17 (1.05, 1.3)	0.0033
Severe steatosis (S1-2 vs S3)	Recessive	Overall	8	0.67	0	1.08 (.78, 1.5)	0.642
Severe steatosis (S1-2 vs S3)	Recessive	Non-Caucasian	1	NA	NA	1.11 (.39, 3.16)	0.852
Severe	Recessive	Caucasian	7	0.72	0	1.08	0.676

steatosis (S1-2 vs S3)						(.76, 1.54)	
NASH (NAFL vs NASH)	Recessive	Overall	9	0.33	0.15	1.14 (.96, 1.36)	0.128
NASH (NAFL vs NASH)	Recessive	Non-Caucasian	3	0	0.58	1.24 (.81, 1.9)	0.324
NASH (NAFL vs NASH)	Recessive	Caucasian	6	0.53	0.06	1.14 (.93, 1.41)	0.213
Any fibrosis (F0 vs F1-4)	Recessive	Overall	9	0.52	0.03	1.27 (1.04, 1.54)	0.0183
Any fibrosis (F0 vs F1-4)	Recessive	Non-Caucasian	2	0	0.82	2.14 (1.2, 3.84)	0.0105
Any fibrosis (F0 vs F1-4)	Recessive	Caucasian	7	0.51	0.06	1.19 (.99, 1.45)	0.068
Advanced fibrosis (F0-2 vs F3-4)	Recessive	Overall	8	0	0.65	1.2 (1.02, 1.42)	0.027
Advanced fibrosis (F0-2 vs	Recessive	Non-Caucasian	2	0	0.64	.96 (.5, 1.85)	0.911

F3-4)							
Advanced fibrosis (F0-2 vs F3-4)	Recessive	Caucasian	6	0	0.5	1.22 (1.03, 1.45)	0.0206
HCC (NAFLD-HCC vs NAFLD no-HCC)	Recessive	Overall	4	0	0.95	1.4 (.99, 1.98)	0.056

Table 1. Summary of results in adults from meta-analyses for dichotomous outcomes. Meta-analyses were performed using random effects with subgroup analysis for Caucasian and Non-Caucasian populations. Additive, recessive, and dominant genetic models were tested for all outcomes. Results using a recessive model of inheritance (CC+CT vs. TT) are shown for all outcomes, except for HCC, where a dominant model (CC vs. CT+TT) is shown. Due to use of three genetic models, critical p-value for effect summary is $p_z<0.017$. Full results (with all genetic models) are in Supplementary Table 5. CI, confidence interval; HCC, hepatocellular carcinoma; NAFL, non-alcoholic fatty liver; NASH, non-alcoholic steatohepatitis; OR, odds ratio.

FIGURE LEGENDS

Fig. 1. The effect of rs641738C>T on liver fat. Data from 29,679916 individuals with CT, or MRI, or MRS liver fat. rs641738C>T positively associated with liver fat in Caucasian populations, where data represents standard deviation change in normalized liver fat per T-allele. CI, confidence interval; UKBB, UK BioBank.

Fig. 2. rs641738C>T is associated with higher odds of diagnosis of NAFLD. Data from 52,17333,263 adults (11,3019,713 cases and 40,87223,550 controls) with radiologically or histologically defined steatosis for presence versus absence of NAFLD. CI, confidence interval; LBC, Liver Biopsy Cohort; OR, odds ratio.

Fig. 3. The effect of rs641738C>T on presence of advanced fibrosis in adult patients with NAFLD. Data from 7,6926,211 adults (1,214828 cases and 6,4785,383 controls) with biopsy-proven NAFLD comparing advanced fibrosis (F3-4) versus F0-2, using a recessive model of inheritance (CC+CT vs. TT). CI, confidence interval; LBC, Liver Biopsy Cohort; OR, odds ratio.

Fig. 4. rs641738C>T is associated with higher odds of NAFLD-HCC. Data from 2,328 adults with NAFLD assessing for the presence versus absence of HCC, using a dominant model of inheritance (CC vs. CT+TT).

Fig. 5. rs641738C>T is positively associated with alanine aminotransferase (ALT) in Caucasian populations in genome-wide association studies (GWAS). Meta-analysis of GWAS summary statistics from 609,794 participants for the association between rs641738C>T on logarithmically-transformed ALT using linear regression. CI, confidence interval; UKBB, UK BioBank.